Physica A 503 (2018) 969–980

Contents lists available at ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

An empirical investigation and theoretic modeling for the collective online visiting behaviors

Li Zhenpeng [a,*], Tang Xijin [b,c], Zhou Haijun [d], Donghui Yan [e]

[a] Department of Mathematics and Computer Sciences, Dali University, Dali, 671003, China
[b] Academy of Mathematics and Systems Sciences Chinese Academy of Sciences, Beijing, 100190, China
[c] University of Chinese Academy of Sciences, Beijing, 100049, China
[d] School of Physical Sciences, University of Chinese Academy of Sciences, Beijing 100049, China
[e] Department of Mathematics, University of Massachusetts, Dartmouth, United States

HIGHLIGHTS

- Exponential novelty decay of replying behavior to posts on BBS results in a linear increase in cumulative clicks at the first several hours, which then gradually saturate to a stable level.
- The amount of cumulative clicks of the posts satisfy Geometric Brownian Motion and log-normal distribution.
- Sensation-stimulus law of psychophysics might explain the correlation between clicks and replies.

ARTICLE INFO

Article history:
Received 6 December 2017
Received in revised form 6 April 2018
Available online xxxx

Keywords:
Sensation-stimulus law
Geometric Brownian motion
Log-normal distribution
Collective novelty decaying
Power law

ABSTRACT

We study Tianya club (one of influential Chinese Bulletin Board System, or BBS), Tianya Zatan board users' clicks and replies behavior. We investigate the underlying statistics of clicks and replies, unravel the mechanism of the collective novelty decay and provide a probabilistic model to describe the users' cumulative clicking patterns. We find that the amount of clicks of the posts satisfy Geometric Brownian Motion and log-normal distribution, both the collective novelty and inter-replying time have an exponential decaying characteristic, and the amount of replies and clicks follow the power law distribution. Finally, sensation-stimulus law of psychophysics is used to explain the correlation between clicks and replies. This paper provides valuable insights for the collective online visiting behaviors, as well as an empirical and theoretic understanding for the massive online attention decaying principle.

1. Introductions

Quantitatively analyzing collective behavior, mining and modeling temporal and spatial patterns is an important focus of statistical physics and complexity sciences. Over the past decades, physicists and sociologists have uncovered many remarkable statistical properties and patterns of collective behaviors, such as the migration and mobility patterns [1,2], online visiting [3], market transaction [4], human communication [5–7], etc.

The general assumption in the previous studies on collective behaviors is the temporal statistics are uniform and stationary, for example, a Poisson process. However, in recent years, the dynamics of online collective behaviors has been shown to possess some regular. For example, the scaling laws of human travel [1], the bimodal distribution in human

* Corresponding author.
E-mail addresses: lizhenpeng@amss.ac.cn (Z. Li), xjtang@amss.ac.cn (X. Tang), zhouhj@itp.ac.cn (H. Zhou), dyan@umassd.edu (D. Yan).

https://doi.org/10.1016/j.physa.2018.08.090
0378-4371/

communication [7], the bursts and heavy-tailed distribution in human communication [5,6,8]. Those non-Poisson properties are found in many types of human behaviors [3,9,10]. These temporal human dynamics have been found to follow a reproducible and predictable pattern, and numerous scientists are delving deep towards understanding the mechanism of collective behavior's [5,11,12].

Thanks to the information technology and the extensive use of social media, a wealth of information and massive data are generated through online social network, for example BBS. This has resulted in a countless number of clicks, comments, replies, updates and retweets of participants. It is possible to investigate individual user behavior, and then by analyzing and aggregating massive user-generated data streams to study collective actions at the population level. Alternatively, large-scale empirical datasets enable the development of detailed computational models of group behavior.

The statistical mechanism of replies and clicks by online users is of great significance for understanding the human dynamics, improving the quality of online services and promoting group behavior guidance and control. Due to the availability of massive datasets, in-depth empirical analysis and the modeling of temporal patterns of online visiting behaviors help in revealing collective dynamics, and explaining many complex socio-economic phenomena. They are also useful in finding applications in online opinions dynamics, recommendation, service allocation and human-behavior prediction, and so on.

[13] provides a comprehensive and state-of-the-art review on progresses in this field, such as the empirical analysis and modeling on human temporal activities, mobility patterns, and the applications of human dynamics.

Empirical analysis and the modeling of temporal patterns of online visiting behaviors is also closely related to dynamics of collective attention, and is at the heart of decision making and the dissemination of ideas. It is thus of interest to understand how attention to novel items propagates and eventually fades among a large population.

In recent years, many studies on the online collective attention use social media consecutive clicking sequence as an empirical evidence to uncover various large scale statistic behavioral phenomena and attention patterns. The subject of collective attention is at the center of digital era. It is arguable to say that, in the era of information explosion, the most valuable resource is attention rather than information.

In general, the online browsing behaviors of users includes clicks and replies. In the previous studies, the click stream is used to study the collective direction of attention diffusion [14], novelty decay [15], or human knowledge maps [16]. However, the studies on clicks behavior rarely involve the relation between clicks and replies. In general for a post in online discussing communities, the collective clicking volume increase may stimulate more active replying behavior, i.e., increasing replying volume. On the other hand, the replying behavior may also affect the users' clicking impulsion, for example, more replying volume for a post will trigger more users' interests to click the post.

The investigations on social media clicking statistics such as click are important gateway to study online collective behaviors. Since the click stream can reveal the massive users' attention fluctuation, information diffusion , topic lifetime in communities, and the time-sensitive preference of users in online surfing. Uncovering and predicting the dynamic process of collective attention have an important theoretical significance on human attention theory, and practical application such as attention marketing, information dissemination.

We note that previous empirical investigations have confirmed that the amounts of views and replies to a post on BBS follow the power law distribution with different power exponent, and the amounts of view and reply of posts have a nonlinear relationship, such properties are shown to be a universal [17]. These results are consistent with our empirical analysis conclusion. However, this study approaches from a mechanism point of view, and use the sensation-stimulus law of psychophysics to explain the nonlinear relationship between clicks and replies.

Our goal here is to uncover and model the user behavior on the Tianya Zatan board, one influential Chinese BBS. We investigate the statistics of replies and clicks, provide a collective novelty exponential decaying mechanism and a probabilistic model to describe the users' cumulative clicking patterns. In particular, based on the collective novelty exponential decaying mechanism, we present rigorous mathematic analysis to interpret the dynamic process of BBS posts cumulative clicks.

2. Data sources and materials

Founded in 1999, as a famous Chinese BBS, Tianya (or Tianya club) is the leading public BBS social media platform. It includes different boards such as public voice, entertainment gossip, emotional world, and Tianya Zatan etc. With respect to popularity, visiting volume and influential power, Tianya Zatan board is one of the leading sections of Tianya. It advocates self-attention, social concern, humanity concern, and plays the role of leading real world public opinion, with rapid sensitive response and distinctive position.

In this paper, we analyze the dataset with 1,241,674 posts published in Tianya Zatan board during the time span from 2012/1/1 00:00:00 to 2015/12/31 23:59:00. The information of each post includes the initial posting time, continuous reply time stamps, cumulative clicking volume, and cumulative replies. For the 1,241,674 posts, we count each post's reply and click volume, around 0.186% of the total posts, or 23,117 posts, have zero reply. After removing these items, we are left with 1,218,557 posts.